<u>RAW SEQUENCE LISTING</u>

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _10/028,245A_

Source: _1Fw/6_

Date Processed by STIC: _5/23/05_

ENTERED

```
    RAW SEQUENCE LISTING                    DATE: 05/23/2005
    PATENT APPLICATION:  US/10/028,245A     TIME: 09:30:02

         Input Set : A:\GC700-SEQLIST2.TXT
         Output Set: N:\CRF4\05232005\J028245A.raw

 4 <110> APPLICANT: Dunn-Coleman, Nigel
 5        Goedegebuur, Frits
 6        Ward, Michael
 7        Yao, Jian
 9 <120> TITLE OF INVENTION: EGVIII Endoglucanase and Nucleic Acids
10        Encoding the Same
12 <130> FILE REFERENCE: GC700
14 <140> CURRENT APPLICATION NUMBER: US 10/028,245A
15 <141> CURRENT FILING DATE: 2001-12-18
17 <160> NUMBER OF SEQ ID NOS: 5
19 <170> SOFTWARE: FastSEQ for Windows Version 4.0
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 1826
23 <212> TYPE: DNA
24 <213> ORGANISM: Trichoderma reesei
26 <400> SEQUENCE: 1
27 gtcgacccac gcgtccgttc attcttcctc ccctcctcct cctcctcctc ctccccttct        60
28 cccccatcac cgtcaccact ctcctcattg ccgctctctc tgcgagccat gacgcagcat       120
29 taacacacac tcgtttctgt tactctcgct gtcgtcggct ctgctcgttg gcattctgct       180
30 tagcactttg tttcgttctt cgttctcttt taatccgtca tcttctgcaa tctgctgcca       240
31 tttgttcgac taggtagtgg taatatacgg acagcttttt ttccctcgct caacacgtcg       300
32 acgtacaatt aatacaccat ctcgttaatc ggatatatcc ctcggcctct tcctggtgct       360
33 tgtgcgacgc tcctcgtttc tccctctcat tatgcgcgca acctcccttc tggccgccgc       420
34 cttggccgtg gctggcgatg ccctcgccgg caagatcaaa tatctgggcg tcgccattcc       480
35 cggaatcgac tttggctgcg acatcgacgg cagctgtccg actgacacgt cgtctgtgcc       540
36 cctgctgagc tacaaaggag gagatggcgc cggccagatg aagcatttcg ccgaagacga       600
37 cggcctcaac gtctttcgca tatccgctac atggcagttt gtcctcaaca cacggtgga       660
38 cggcaagctg gacgagctca actggggctc ctacaacaag gtcgtcaacg cctgtctcga       720
39 gacgggcgcc tactgcatga ttgacatgca caactttgcc cgctacaacg cggcatcat       780
40 cggccaggga ggcgtgtcgg acgacatctt tgtcgacctc tgggtccaga tcgcaaagta       840
41 ctacgaggac aacgacaaga tcatctttgg cctgatgaac gagccgcacg acctcgacat       900
42 tgagatctgg gcgcagacgt gccaaaaggt cgtcactgcg atccgaaagg ccggcgccac       960
43 ctcgcagatg atcctcctgc ccggaaccaa ctttgccagc gtcgagacgt atgtgtccac      1020
44 tggcagcgcg gaagccctcg gcaagattac gaacccggat ggaagcaccg atttgctgta      1080
45 ctttgatgtc cacaagtatc tcgacatcaa caactccggg tcgcacgccg agtgcaccac      1140
46 agacaacgtc gacgccttca acgacttcgc ggactggctg aggcagaaca agcgccaggc      1200
47 catcatctcc gaaacgggcg cgtccatgga accttcgtgc atgactgcct tctgcgccca      1260
48 gaacaaggcc attagcgaaa acagcgacgt ctacattggc tttgtgggct ggggtgccgg      1320
49 cagctttgac acgtcgtaca tcttgactct gactcccctc ggcaagcccg gcaactacac      1380
50 cgacaacaag ctcatgaacg agtgcattct ggaccagttt accctcgacg aaaagtaccg      1440
51 tccaacaccc acctcaattt ccacagcggc ggaagagacg ccacggcga cagcaacctc      1500
52 tgacggcgac gcgccatcca ctacgaagcc catctttagg gaagaaaccg cctctcccac      1560
53 tcccaatgct gttaccaagc cctcgcccga cacgagcgac tcttccgacg acgacaagga      1620
```

RAW SEQUENCE LISTING   DATE: 05/23/2005
PATENT APPLICATION: US/10/028,245A   TIME: 09:30:02

Input Set : A:\GC700-SEQLIST2.TXT
Output Set: N:\CRF4\05232005\J028245A.raw

```
54 ctcggcagca tctatgagtg cccagggctt gacaggcacg gtgctgttta ctgttgctgc   1680
55 ccttggctac atgctggtag cgttttgatg ttttttttt aatgagtttg tatacctaat   1740
56 gagcatgatt gagatgctac gtagtatata tgtctttacg ggtacataag actagagcca   1800
57 tgttgtaatc aaaaaaaaaa aaaaaa                                         1826
59 <210> SEQ ID NO: 2
60 <211> LENGTH: 419
61 <212> TYPE: PRT
62 <213> ORGANISM: Trichoderma reesei
64 <400> SEQUENCE: 2
65 Gly Lys Ile Lys Tyr Leu Gly Val Ala Ile Pro Gly Ile Asp Phe Gly
66  1               5                  10                  15
67 Cys Asp Ile Asp Gly Ser Cys Pro Thr Asp Thr Ser Ser Val Pro Leu
68              20                  25                  30
69 Leu Ser Tyr Lys Gly Gly Asp Gly Ala Gly Gln Met Lys His Phe Ala
70          35                  40                  45
71 Glu Asp Asp Gly Leu Asn Val Phe Arg Ile Ser Ala Thr Trp Gln Phe
72      50                  55                  60
73 Val Leu Asn Asn Thr Val Asp Gly Lys Leu Asp Glu Leu Asn Trp Gly
74 65                  70                  75                  80
75 Ser Tyr Asn Lys Val Val Asn Ala Cys Leu Glu Thr Gly Ala Tyr Cys
76                  85                  90                  95
77 Met Ile Asp Met His Asn Phe Ala Arg Tyr Asn Gly Gly Ile Ile Gly
78              100                 105                 110
79 Gln Gly Gly Val Ser Asp Asp Ile Phe Val Asp Leu Trp Val Gln Ile
80          115                 120                 125
81 Ala Lys Tyr Tyr Glu Asp Asn Asp Lys Ile Ile Phe Gly Leu Met Asn
82      130                 135                 140
83 Glu Pro His Asp Leu Asp Ile Glu Ile Trp Ala Gln Thr Cys Gln Lys
84 145                 150                 155                 160
85 Val Val Thr Ala Ile Arg Lys Ala Gly Ala Thr Ser Gln Met Ile Leu
86              165                 170                 175
87 Leu Pro Gly Thr Asn Phe Ala Ser Val Glu Thr Tyr Val Ser Thr Gly
88              180                 185                 190
89 Ser Ala Glu Ala Leu Gly Lys Ile Thr Asn Pro Asp Gly Ser Thr Asp
90          195                 200                 205
91 Leu Leu Tyr Phe Asp Val His Lys Tyr Leu Asp Ile Asn Asn Ser Gly
92      210                 215                 220
93 Ser His Ala Glu Cys Thr Thr Asp Asn Val Asp Ala Phe Asn Asp Phe
94 225                 230                 235                 240
95 Ala Asp Trp Leu Arg Gln Asn Lys Arg Gln Ala Ile Ile Ser Glu Thr
96              245                 250                 255
97 Gly Ala Ser Met Glu Pro Ser Cys Met Thr Ala Phe Cys Ala Gln Asn
98              260                 265                 270
99 Lys Ala Ile Ser Glu Asn Ser Asp Val Tyr Ile Gly Phe Val Gly Trp
100         275                 280                 285
101 Gly Ala Gly Ser Phe Asp Thr Ser Tyr Ile Leu Thr Leu Thr Pro Leu
102     290                 295                 300
103 Gly Lys Pro Gly Asn Tyr Thr Asp Asn Lys Leu Met Asn Glu Cys Ile
104 305                 310                 315                 320
```

```
105 Leu Asp Gln Phe Thr Leu Asp Glu Lys Tyr Arg Pro Thr Pro Thr Ser
106                 325             330             335
107 Ile Ser Thr Ala Ala Glu Glu Thr Ala Thr Ala Thr Ala Thr Ser Asp
108                 340             345             350
109 Gly Asp Ala Pro Ser Thr Thr Lys Pro Ile Phe Arg Glu Glu Thr Ala
110             355             360             365
111 Ser Pro Thr Pro Asn Ala Val Thr Lys Pro Ser Pro Asp Thr Ser Asp
112         370             375             380
113 Ser Ser Asp Asp Asp Lys Asp Ser Ala Ala Ser Met Ser Ala Gln Gly
114 385             390             395             400
115 Leu Thr Gly Thr Val Leu Phe Thr Val Ala Ala Leu Gly Tyr Met Leu
116             405             410             415
117 Val Ala Phe
119 <210> SEQ ID NO: 3
120 <211> LENGTH: 19
121 <212> TYPE: PRT
122 <213> ORGANISM: Trichoderma reesei
124 <400> SEQUENCE: 3
125 Met Arg Ala Thr Ser Leu Leu Ala Ala Ala Leu Ala Val Ala Gly Asp
126  1               5              10              15
127 Ala Leu Ala
129 <210> SEQ ID NO: 4
130 <211> LENGTH: 1317
131 <212> TYPE: DNA
132 <213> ORGANISM: Trichoderma reesei
134 <400> SEQUENCE: 4
135 atgcgcgcaa cctcccttct ggccgccgcc ttggccgtgg ctggcgatgc cctcgccggc      60
136 aagatcaaat atctgggcgt cgccattccc ggaatcgact ttggctgcga catcgacggc     120
137 agctgtccga ctgacacgtc gtctgtgccc ctgctgagct acaaaggagg agatggcgcc     180
138 ggccagatga agcatttcgc cgaagacgac ggcctcaacg tctttcgcat atccgctaca     240
139 tggcagtttg tcctcaacaa cacggtggac ggcaagctgg acgagctcaa ctggggctcc     300
140 tacaacaagg tcgtcaacgc ctgtctcgag acgggcgcct actgcatgat tgacatgcac     360
141 aactttgccc gctacaacgg cggcatcatc ggccagggag gcgtgtcgga cgacatcttt     420
142 gtcgacctct gggtccagat cgcaaagtac tacgaggaca acgacaagat catctttggc     480
143 ctgatgaacg agccgcacga cctcgacatt gagatctggg cgcagacgtg ccaaaaggtc     540
144 gtcactgcga tccgaaaggc cggcgccacc tcgcagatga tcctcctgcc cggaaccaac     600
145 tttgccagcg tcgagacgta tgtgtccact ggcagcgcgg aagccctcgg caagattacg     660
146 aacccggatg gaagcaccga tttgctgtac tttgatgtcc acaagtatct cgacatcaac     720
147 aactccgggt cgcacgccga gtgcaccaca gacaacgtcg acgccttcaa cgacttcgcg     780
148 gactggctga ggcagaacaa gcgccaggcc atcatctccg aaacgggcgc gtccatggaa     840
149 ccttcgtgca tgactgcctt ctgcgcccag aacaaggcca ttagcgaaaa cagcgacgtc     900
150 tacattggct ttgtgggctg gggtgccggc agctttgaca cgtcgtacat cttgactctg     960
151 actcccctcg gcaagcccgg caactacacc gacaacaagc tcatgaacga gtgcattctg    1020
152 gaccagttta ccctcgacga aaagtaccgt ccaacaccca cctcaatttc cacagcggcg    1080
153 gaagagacgg ccacggcgac agcaacctct gacggcgacg cgccatccac tacgaagccc    1140
154 atctttaggg aagaaaccgc ctctcccact cccaatgctg ttaccaagcc ctcgcccgac    1200
155 acgagcgact cttccgacga cgacaaggac tcggcagcat ctatgagtgc ccagggcttg    1260
156 acaggcacgg tgctgtttac tgttgctgcc cttggctaca tgctggtagc gttttga       1317
158 <210> SEQ ID NO: 5
```

159 <211> LENGTH: 438
160 <212> TYPE: PRT
161 <213> ORGANISM: Trichoderma reesei
163 <400> SEQUENCE: 5
164 Met Arg Ala Thr Ser Leu Leu Ala Ala Ala Leu Ala Val Ala Gly Asp
165 1   5   10   15
166 Ala Leu Ala Gly Lys Ile Lys Tyr Leu Gly Val Ala Ile Pro Gly Ile
167   20   25   30
168 Asp Phe Gly Cys Asp Ile Asp Gly Ser Cys Pro Thr Asp Thr Ser Ser
169   35   40   45
170 Val Pro Leu Leu Ser Tyr Lys Gly Gly Asp Gly Ala Gly Gln Met Lys
171   50   55   60
172 His Phe Ala Glu Asp Asp Gly Leu Asn Val Phe Arg Ile Ser Ala Thr
173 65   70   75   80
174 Trp Gln Phe Val Leu Asn Asn Thr Val Asp Gly Lys Leu Asp Glu Leu
175   85   90   95
176 Asn Trp Gly Ser Tyr Asn Lys Val Val Asn Ala Cys Leu Glu Thr Gly
177   100   105   110
178 Ala Tyr Cys Met Ile Asp Met His Asn Phe Ala Arg Tyr Asn Gly Gly
179   115   120   125
180 Ile Ile Gly Gln Gly Gly Val Ser Asp Asp Ile Phe Val Asp Leu Trp
181   130   135   140
182 Val Gln Ile Ala Lys Tyr Tyr Glu Asp Asn Asp Lys Ile Ile Phe Gly
183 145   150   155   160
184 Leu Met Asn Glu Pro His Asp Leu Asp Ile Glu Ile Trp Ala Gln Thr
185   165   170   175
186 Cys Gln Lys Val Val Thr Ala Ile Arg Lys Ala Gly Ala Thr Ser Gln
187   180   185   190
188 Met Ile Leu Leu Pro Gly Thr Asn Phe Ala Ser Val Glu Thr Tyr Val
189   195   200   205
190 Ser Thr Gly Ser Ala Glu Ala Leu Gly Lys Ile Thr Asn Pro Asp Gly
191   210   215   220
192 Ser Thr Asp Leu Leu Tyr Phe Asp Val His Lys Tyr Leu Asp Ile Asn
193 225   230   235   240
194 Asn Ser Gly Ser His Ala Glu Cys Thr Thr Asp Asn Val Asp Ala Phe
195   245   250   255
196 Asn Asp Phe Ala Asp Trp Leu Arg Gln Asn Lys Arg Gln Ala Ile Ile
197   260   265   270
198 Ser Glu Thr Gly Ala Ser Met Glu Pro Ser Cys Met Thr Ala Phe Cys
199   275   280   285
200 Ala Gln Asn Lys Ala Ile Ser Glu Asn Ser Asp Val Tyr Ile Gly Phe
201   290   295   300
202 Val Gly Trp Gly Ala Gly Ser Phe Asp Thr Ser Tyr Ile Leu Thr Leu
203 305   310   315   320·
204 Thr Pro Leu Gly Lys Pro Gly Asn Tyr Thr Asp Asn Lys Leu Met Asn
205   325   330   335
206 Glu Cys Ile Leu Asp Gln Phe Thr Leu Asp Glu Lys Tyr Arg Pro Thr
207   340   345   350
208 Pro Thr Ser Ile Ser Thr Ala Ala Glu Glu Thr Ala Thr Ala Thr Ala

RAW SEQUENCE LISTING   DATE: 05/23/2005
PATENT APPLICATION: **US/10/028,245A**   TIME: 09:30:02

Input Set : **A:\GC700-SEQLIST2.TXT**
Output Set: **N:\CRF4\05232005\J028245A.raw**

```
209         355                    360                    365
210 Thr Ser Asp Gly Asp Ala Pro Ser Thr Thr Lys Pro Ile Phe Arg Glu
211     370                    375                    380
212 Glu Thr Ala Ser Pro Thr Pro Asn Ala Val Thr Lys Pro Ser Pro Asp
213 385                    390                    395                    400
214 Thr Ser Asp Ser Ser Asp Asp Asp Lys Asp Ser Ala Ala Ser Met Ser
215                    405                    410                    415
216 Ala Gln Gly Leu Thr Gly Thr Val Leu Phe Thr Val Ala Ala Leu Gly
217             420                    425                    430
218 Tyr Met Leu Val Ala Phe
219         435
```

VERIFICATION SUMMARY   DATE: 05/23/2005
PATENT APPLICATION: **US/10/028,245A**   TIME: 09:30:03

Input Set : **A:\GC700-SEQLIST2.TXT**
Output Set: **N:\CRF4\05232005\J028245A.raw**